

## Portland State University PDXScholar

---

Anthropology Faculty Publications and  
Presentations

Anthropology

---

1-2011

# Can Salmonids (*Oncorhynchus* spp.) Be Identified to Species Using Vertebral Morphometrics?

Harriet R. Huber

Jeffery C. Jorgensen

Virginia L. Butler

*Portland State University*, [virginia@pdx.edu](mailto:virginia@pdx.edu)

Greg Baker


*Portland State University*

Rebecca Stevens

*Eastern Washington University*

Let us know how access to this document benefits you.

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/anth\\_fac](https://pdxscholar.library.pdx.edu/anth_fac)

 Part of the [Biological and Physical Anthropology Commons](#), and the [Natural Resources and Conservation Commons](#)

---

### Citation Details

Huber, Harriet R.; Jorgensen, Jeffery C.; Butler, Virginia L.; Baker, Greg; and Stevens, Rebecca, "Can Salmonids (*Oncorhynchus* spp.) Be Identified to Species Using Vertebral Morphometrics?" (2011). *Anthropology Faculty Publications and Presentations*. 18.  
[https://pdxscholar.library.pdx.edu/anth\\_fac/18](https://pdxscholar.library.pdx.edu/anth_fac/18)

This Post-Print is brought to you for free and open access. It has been accepted for inclusion in Anthropology Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

**Can salmonids (*Oncorhynchus* spp.) be identified to species using vertebral morphometrics?**

Harriet R. Huber<sup>a</sup>, Jeffrey C. Jorgensen<sup>b</sup>, Virginia L. Butler<sup>c</sup>, Greg Baker<sup>c</sup>, Rebecca Stevens<sup>d</sup>

<sup>a</sup>National Marine Mammal Laboratory  
National Marine Fisheries Service/Alaska Fisheries Science Center  
7600 Sand Point Way NE  
Seattle, WA 98115

corresponding author: [harriet.huber@noaa.gov](mailto:harriet.huber@noaa.gov)  
phone: 206 526-6433; fax: 206 526-6615

<sup>b</sup> Conservation Biology Division  
National Marine Fisheries Service/Northwest Fisheries Science Center  
2725 Montlake Blvd E.  
Seattle, WA 98112

<sup>c</sup>Department of Anthropology  
Portland State University  
Portland, OR 97207

<sup>d</sup>Archaeological and Historical Services  
Eastern Washington University  
Cheney, WA 99004

Published in *Journal of Archaeological Science* in January, 2011:

Huber, H.R., Jorgensen, J.C., Butler, V.L., Baker, G., and Stevens, R. 2011. Can salmonids (*Oncorhynchus* spp.) be identified to species using vertebral morphometrics? *Journal of Archaeological Science* 38:136-146. doi:10.1016/j.jas.2010.08.020

Salmonid classification models are available on the web:  
<http://www.conserver.iugo-cafe.org>

## ABSTRACT

Remains of anadromous Pacific salmon and trout (genus *Oncorhynchus*) are common in archaeological sites from California to Alaska; however, morphological similarity generally precludes species identification, limiting the range of questions that salmonid remains can address in relation to past human use and ongoing efforts in conservation biology. We developed a relatively simple, rapid, and non-destructive way to classify salmon and trout vertebrae from archaeological contexts to species using length, height and the ratio of length to height. Modern reference material was obtained from all seven anadromous *Oncorhynchus* species native to the west coast of North America. A minimum of ten adult Chinook (*Oncorhynchus tshawytscha*), chum (*O. keta*), coho (*O. kisutch*), pink (*O. gorbuscha*), and sockeye salmon (*O. nerka*) and cutthroat (*O. clarki clarki*) and steelhead trout (*O. mykiss*) were skeletonized and vertebra length and height were measured. Morphometric analyses compared species classification success based on Linear Discriminant Analysis (LDA), Classification and Regression Trees (CART), and randomForest, with CART performing the best. Classification analyses used all seven species individually, but because of considerable overlap among several species we also conducted analyses on four species groupings. We assigned Chinook salmon and cutthroat to their own groups based on their dissimilarities from each other and the other species. The remaining species were divided into two group complexes (a) chum, coho, and steelhead; and (b) pink and sockeye. When we grouped species according to similar morphology, CART overall success rates increased, ranging from 92 to 100%. Individual species with the highest successful classification rates using CART were Chinook salmon and cutthroat, from 92 to 100%, respectively. We applied our classification to an assemblage of

ancient (1000 to 3000 year old) salmonid vertebrae from the Swiftwater Rockshelters excavations on the upper Wenatchee River in Washington State, U.S.A.

Keywords: salmonid morphometrics, vertebrae, Classification and Regression Trees, CART

## 1. Introduction

Hundreds of archaeological sites from Alaska to California contain the remains of salmon and trout (Family Salmonidae), mainly in the genus *Oncorhynchus* (Cannon 2000; Butler and Campbell 2004; Gobalet et al. 2004). Such fish bones have tremendous potential to contribute to a range of issues in archaeology and fisheries science. In the eastern Pacific, *Oncorhynchus* is represented by seven species, which vary greatly in size, abundance, seasonal availability, nutritional value, spawning habitat and other features that greatly affect human use patterns today and, undoubtedly, in the past.

However, traditional faunal analysis relying on skeletal morphology rarely can distinguish species within the genus, thus constraining our ability to understand details of human-fish interactions. Only a few cranial elements are species-diagnostic (Casteel 1974; Gorshkov et al. 1979) and these are recovered infrequently from archaeological sites (probably because of preservation factors, Butler and Chatters 1994). Postcranial elements, mainly vertebrae, dominate archaeological assemblages and while these typically can be recognized as *Oncorhynchus*, distinguishing all species based on surface morphology is challenging (Gobalet et al. 2004). Lumping all the salmon and trout remains into one category keeps us from a detailed understanding of past human land use and subsistence patterns. Moreover, coarse-level identifications limit our ability to study past species distributions, critical to ongoing fish recovery efforts (e.g., Adams et al. 2007; Butler et al. unpublished manuscript).

Breakthroughs in ancient salmon DNA analyses have gone a long way towards addressing this “identification problem”. Yang et al. (2004) have successfully extracted salmon mtDNA from remains taken from multiple archaeological sites in the Pacific Northwest.

Scholars have used the species identifications to address such questions as whether fish were stored, or if species use was mediated by social rank (Speller et al. 2005; Cannon and Yang 2006). While ancient DNA is a powerful new tool, it has its limitations: the method is destructive and expensive, and consequently only a small sample of remains from a given context is typically studied. Researchers have also turned to metrics, particularly vertebra diameter and radiographic analysis of growth rings, to determine salmon species (Cannon and Yang 2006; Orchard and Szpak pers. comm.). While radiographic analysis as a tool for species identification initially showed some promise (Cannon 1988), its limitations are now well established (Cannon and Yang 2006).

Here we describe our effort to develop a rapid and non-destructive way to classify salmon and trout remains from archaeological contexts using morphometric analysis of vertebra shape to classify specimens to species. Previously, Butler and Baker (2003) suggested that vertebra shape, in particular the ratio of height-to-length, was distinctive across species or species groups. This study greatly expands that effort, including a much larger number of reference skeletons and more powerful classification methods. Our goal for this research was to evaluate and develop simple but practical classification tools for investigators, and to inspire further quantitative research and discussion in *Oncorhynchus* species identification via vertebral morphology. We tested the usefulness of three distinct classification methods, Linear Discriminant Analysis (LDA), Classification and Regression Trees (CART), and randomForest, and found that CART provides the overall best results. We applied the CART classification model to the well-preserved salmonid remains from the Swiftwater Rockshelters archaeological site on the Wenatchee River in east-central Washington State, U.S.A.

## 2. Material and methods

### 2.1 Reference collection preparation and measurements

Seven species of anadromous salmon and trout are found in the northeastern Pacific Ocean: Chinook (*Oncorhynchus tshawytscha*), chum (*O. keta*), coho (*O. kisutch*), pink (*O. gorbuscha*), and sockeye (*O. nerka*) salmon; and steelhead (*O. mykiss*) and cutthroat (*O. clarki clarki*) trout. A minimum of ten adults from each species was collected between 1999 and 2000 from various locations throughout Washington State (Table 1). All were of spawning age, and represented body sizes typical of adult spawners. Skeletons were prepared using dermestid beetles so that vertebra morphology was not altered. Spinal columns were kept intact during processing; vertebrae were strung on a wire sequentially to preserve their original position.

Vertebra morphology varies along the column (as described by Morales 1984; Butler 1990, 1993). We focused our classification on two of the four types of vertebrae (after Butler 1993), Types II and III make up over 90% of the column. Following Butler (1993), Type II are those with unfused neural and haemal processes (Fig. 1; A1, A2, A3), Type III have fused processes (Fig. 1; B). We based our classification models on vertebra length, height, and the ratio between the two because these measures had shown potential value in previous studies and can be relatively quickly and consistently measured. As shown in Figure 2A, height refers to the height of the centrum face from the dorsal to ventral margins of the centrum. Vertebral length is the distance along the lateral margin from the edge of the centrum on the rostral face to the centrum edge on the caudal face (Fig. 2B). Length and height were measured three times and the means of the three measurements were used for analysis. Multiple measures of the same element allowed us to identify measurement error and degree of replicability. Two of us (HRH,

JCJ) measured length and height using electronic calipers (Mitutoyo Digimatic<sup>1</sup>) to the nearest 0.01 mm. We randomly selected 30 bones and tested for differences in measurements between observers using a paired t-test and found no significant difference ( $t = 1.94$ ,  $df = 29$ ,  $p > 0.05$ ) in mean length measurements.

In archaeological studies, other measures of vertebra size have been defined such as width (Casteel 1976, see Figure 46; Cannon 1988) or transverse diameter (Cannon and Yang 2006). These refer to the measured distance 90 degrees from height as used in this study.

A total of 4,463 Type II and III modern vertebrae were included in our analysis (Table 2). All of our vertebral measurements are available by request to [jeff.jorgensen@noaa.gov](mailto:jeff.jorgensen@noaa.gov).

## *2.2 Classification techniques*

We evaluated the ability of three powerful classification methods to separate vertebrae according to species or species groupings. First, we employed Linear Discriminant Analysis (LDA), a common classification technique. Then, because bone morphology was similar between several species, making them difficult to separate, we applied two tree-based methods: Classification and Regression Trees (CART) and randomForest. We chose tree-based methods because of their ease of use, interpretability, and because all characteristics are considered when determining each branching split. Further, tree-based methods can resolve complex interactions between characteristics that may not be apparent using linear methods. Tree-based classification analysis is being employed more frequently in archaeology in recent years (Feldesman 2002; Weinand 2007).

---

<sup>1</sup> Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA



The entire data set ( $n=4463$ ) comprised a training set and a test set. The training set ( $n=4,393$ ) was used to determine classification criteria and the test set ( $n=70$ ) was used to evaluate the success rates of the classification methods. The test set was made up of ten randomly selected vertebrae from each species (five of each vertebra type). We used open access software R for all data analyses (R Development Core Team 2008; <http://www.r-project.org/>).

### *2.2.1 Linear discriminant analysis (LDA)*

In this widely employed technique, class assignment occurs by fitting linear discriminant functions, derived from the descriptive variables, positioned to bisect the data, which minimizes residuals from the discriminant functions (e.g., Ripley 1996). The data are assumed to be normal and missing values are not permitted. We used the LDA implementation in R (R Development Core Team 2008) in the MASS package (Venables and Ripley 2002).

### *2.2.2 Classification and regression trees (CART)*

This tree-based method works on the principle of splitting or partitioning the data along left and right binary branchings in a recursive manner that minimizes node impurity (Breiman et al. 1984). It has been applied in many fields and has been described in detail elsewhere (see Feldesman 2002 for an extensive introduction). All characteristics are evaluated at each branching of the data which splits observations into nodes (i.e., if length <10.5 mm then left branch, else right branch). CART uses a splitting rule that chooses the one characteristic that minimizes node impurity and improves classification accuracy. Node impurity increases as node

diversity increases i.e., as the number of classes in a node increases. Nodes are split further until observations are placed into terminal nodes. In terminal nodes no further splitting occurs.

Terminal nodes typically consist of observations all belonging to the same class. A tree could be grown so that each observation occupies its own terminal node, but such a tree would have little predictive power on new observations since it overfits the data. Thus, a tree is grown and then “pruned” as a way to prevent overfitting. Terminal nodes are pruned to maximize tree complexity (measured as the number of terminal nodes) and also to minimize the rate of misclassification on new independent data (Feldesman 2002; Karels et al. 2004). CART requires no prior assumptions about the structure of the data, thus normality requirements for parametric methods such as LDA can be relaxed. The classification models we fitted used all three characteristics, length, height, and the length/height ratio. Including all three characteristics was important for the models to discern specimens to species. Furthermore, potentially correlated variables aren’t an issue for CART (Breiman et al. 1984; Karels et al. 2004).

A short outline of our procedure follows. A tree was grown and as each new terminal node was added the error rate was calculated using ten-fold cross-validation. In ten-fold cross-validation, the data were randomly partitioned into ten equal parts, where nine of the ten parts were used to fit a tree model and one part of the data was used to calculate the tree error rate. The tree with the lowest error rate was chosen and that tree was grown with all the data. We used the 1-SE rule to prune the tree; we removed terminal nodes to the point where the tree’s misclassification cost was within  $\pm 1$  SE of the cross-validation error rate (Therneau and Atkinson 1997; Karels et al. 2004). In this way, the misclassification cost was minimized to protect against overfitting. We used the R package rpart (R Development Core Team 2008; Therneau and Atkinson 2008).

### *2.2.3 Random forest*

Also a tree-based method, randomForest is an ensemble or forest of many tree classifiers, with the class determined by a majority vote of the forest of trees (Breiman 2001; see Cutler et al. (2007) for a recent and detailed ecological application). In contrast to CART, this technique has two random components. First, a bootstrap sample of the data is generated (with replacement), where approximately one-third of the data is left “out-of-bag” to estimate the classification error rate, and a tree is grown. This procedure is repeated many times and a simple majority of votes, from all of the trees grown for a particular class, is used to assign classes for observations. Second, branching splits are made as in CART; however, a subset from the full set of predictors is randomly chosen and then evaluated at each node. Like CART, the variable among the subset that minimizes node impurity is chosen for the branching; but, in contrast to CART, trees are not pruned. These two procedures combined tend to increase classification accuracy and reduce the potential for overfitting. The out-of-bag classification error rate has been shown to very closely approximate the error rate on test data (Breiman 2001). The implementation in R (Liaw and Weiner 2002) appears to be fairly insensitive to the number of trees grown; we chose the default of 500 trees.

### *2.3 Archaeological site and sample analysis*

We obtained *Oncorhynchus* spp. vertebrae from a faunal collection excavated from Swiftwater Rockshelters (45CH433), a middle to late period multiple component site located approximately 20 river miles downstream of Lake Wenatchee on the Wenatchee River, a

tributary of the Columbia River in east-central Washington State (Fig. 3). Based on the dominance of salmonid remains in the faunal assemblage, this site is interpreted as a seasonal fishing encampment although other subsistence activities are indicated as well (Lyman 2003). Marine shell beads and Oregon obsidian suggest that inhabitants participated in far-reaching trade networks (Stevens 2003). Three main strata were identified. Radiocarbon ages place human occupation in the upper stratum to be 1040 years before present (ybp) and the lower stratum ranged from 2420 to 2900 ybp ( $\pm 60$  yr). The two strata are separated by a deposit of fluvial sands (Stevens 2003). The faunal remains were extremely well-preserved. Two *Oncorhynchus* cranial elements were sufficiently complete to identify to species using morphology: an otolith from Chinook salmon and a lingual plate from sockeye salmon (Butler and Baker 2003). Two sets of articulated vertebrae ( $n=11$ ,  $n=24$ ; representing 8.5% of the total vertebrae excavated from the site), each belonging to an individual fish, were found in the lower stratum. We included them in overall analyses but excluded all but one from each set in the comparisons of species composition between strata to avoid overinflating the count of that species.

We measured height and length of all the vertebra Type II and III in site deposits with intact margins, according to the above protocol. This resulted in the measurement of 365 vertebrae. We applied the classification models in three ways. First, we applied the models developed with all seven salmonid species. Second, we refitted the classification models, excluding two *Oncorhynchus* species (chum, pink) that probably did not occur in the Wenatchee River basin, based on 19th and 20th century records (Mullan et al. 1992; Hard et al. 1996; Johnson et al 1997). Finally, we refitted the classification models using our four groupings determined by similarities in vertebral morphology: Chinook salmon; steelhead trout, chum and

coho salmon; pink and sockeye salmon; and cutthroat trout. The classification models were fitted to the reference collection for each vertebrae type (II and III) separately, and we assigned excavated vertebrae to species using the models corresponding to their vertebral types.

### **3. RESULTS**

#### *3.1 Modern reference collection*

The three classification methods assigned training set vertebrae to the seven salmonid species with varying success. Vertebrae from Chinook salmon, because of their large size, and cutthroat trout, because of their small size, had the highest classification success rates of the seven species. CART had the best performance among the three methods as measured by the overall success rates followed by randomForest and LDA (Table 3). Among species, cutthroat were the most successfully classified by all the methods tested (range: 0.91 – 1.0), followed by Chinook (range: 0.87 – 0.96). Chum had the most variable and lowest successful classification rates (range: 0.14 – 0.73). Using CART, the rate of successful classification was higher using Type III vertebrae for all species except coho and chum. For nearly all species, success rates of the classification methods were lower when we combined both Types II and III vertebrae in the classification procedure. When we applied the test data set to the classification procedures developed from the training set, CART was still the best classifier, followed by randomForest and then LDA (Table 4). Once again, Chinook and cutthroat were most easily distinguished and chum, coho and steelhead the most likely to be misclassified.

Given the considerable overlap in height, length and length/height ratio measurements among five of the species, and our desire for a coarse tool to differentiate *Oncorhynchus*

vertebrae, we aggregated these species into two group complexes: (a) chum, coho, and steelhead; and (b) pink and sockeye. Chinook and cutthroat remained as separate classes given their distinctiveness. Fig. 4 shows the degree of overlap in height measurements among the five species in the two group complexes.

Classification accuracy increased using these groupings. CART remained the best classification technique with the highest success rates (range: 0.92 – 0.96; Table 5). Overall success rates ranged from 0.89 – 0.96. Cutthroat were consistently classified with the highest success (range: 0.92 – 1.0). The group consisting of chum, coho, and steelhead had the highest variability in success rate (range: 0.67 – 0.95) using all three techniques, but the CART technique had the highest success rate (0.91-0.95). Using Type III vertebrae, CART was best at distinguishing the four groups with the training set (Table 5) and CART and randomForest were equally good at distinguishing the four groups with the test set (Table 6). There were no clear patterns in the success rate when we compared vertebral type isolates (either Type II or III, or II and III combined). The test vertebrae set had a dramatically higher class prediction success rate with the species groupings (Table 6) compared to predictions made for all species discretely (Table 4). Successful classification rates ranged from 0.80 – 1.0. All Chinook salmon vertebrae were classified correctly by all three methods. CART and randomForest predicted classes equally well, and their prediction rates were slightly better than LDA.

### *3.2 Archaeological samples*

Given CART's overall accuracy in making correct species assignments with our reference collection, we focus the reporting of our results using this technique.

When we included all seven species in the mix, sockeye and Chinook dominated the Rockshelters site, followed by pink and steelhead (Table 7). When we removed chum and pink from the classification model, most of the vertebrae assigned to pinks were reclassified as sockeye, with some as coho and a few as steelhead; the single vertebra assigned to chum was reassigned as steelhead; 20% of the bones originally classified as steelhead were reassigned (1 as coho and 6 as sockeye). In the reduced species set, sockeye still dominated the site as a whole; Chinook was second, followed by steelhead and then coho (Table 7).

Species assignment of the two set of articulated vertebral columns in the lower stratum was consistent. All of the vertebrae from each set ( $n = 11$ ;  $n = 24$ ) were assigned to Chinook, using all three methods (LDA, CART, randomForest).

Species abundance varied considerably across strata. With the full species set, Chinook dominated the lower stratum, followed by sockeye and pink (Table 7, Fig. 5). In contrast, sockeye dominated the upper stratum, followed by pink, steelhead, and Chinook (Table 7, Fig. 5).

In the reduced species set with chum and pink removed, most of the vertebrae assigned to pink were reclassified as sockeye (72%), a smaller proportion were reclassified as coho (18%) or steelhead (10%). Sockeye dominated the upper stratum followed by steelhead and Chinook (Table 7, Fig. 5). Chinook and sockeye were evenly distributed in the lower stratum (Table 7). Cutthroat was found only in the upper stratum. Coho was present in the upper and lower strata in small numbers (Table 7, Fig. 5).

#### **4. Discussion**

#### *4.1 Reference collection*

Despite the morphological measurement overlaps, these classification tools were able to assign vertebrae to species with remarkable accuracy especially when we combined several species into groups. CART had the best overall classification prediction performance for most species or groups of species, and the tree-based methods performed better than LDA overall. In general, species assignment improved when vertebral type was known, and substantially improved for the species groups.

Besides CART's implementation in commercial statistical software packages (e.g., Matlab, SPlus), it is readily accessible in open source software. These methods are relatively easy to set up and use. All of these methods can incorporate prior information about the likelihood of certain species being present before assigning new vertebrae to species--i.e., species classes could be up- or down-weighted prior to making predictions on test data. Further, these methods are capable of providing not only the species assignment of each of the vertebrae, but also the probabilities of each vertebra belonging to each of the seven salmonid species. In our case we assigned vertebrae to species based on the highest probability (either a posterior or the counts of a majority vote); several of these procedures could be combined as a weight of evidence to designate either the species or species group. In addition, the tree-based methods (CART, randomForest) are more flexible than LDA, they aren't dependent on data normality assumptions and they allow missing values.

Tree-based methods are a departure from the more widely used discriminant analyses (Feldesman 2002) and their use and interpretation are somewhat different; however, overall they performed better (CART, in particular) than LDA with our data, and classification performance



varied somewhat within species, among species groups, and across the vertebral types. For example, under the CART method both sockeye and steelhead Type III vertebrae were more often correctly classified to their corresponding species (in both the training and test data) than Type II vertebrae or by the combination of these types (Table 3). Vertebrae were assigned to groups with success rates of 0.90 and above regardless of vertebrae type. Although not explored here, classification models could be fitted to the species comprising just the groups (chum, coho, and steelhead; pink and sockeye) as a kind of post hoc classification procedure that could be used to resolve to species those test set vertebrae that were assigned to these groups.

Given the overlap in vertebral measures we observed, one possible concern in our study that relies on vertebral shape (as described by length, height, and length/height) to differentiate species, is could fishes of the same size, regardless of species, have the same vertebral shape? Our results show that fishes of different species that are similar in total length (all collected during their spawning life stage) have considerable differences in vertebral shapes. For example, the fish length ranges of chum, coho, steelhead, and Chinook in our reference samples all overlap each other; however, very few of the Chinook salmon vertebrae were misclassified into the chum, coho, or steelhead complex.

#### *4.2 Archaeological samples*

For the Swiftwater Rockshelters site overall, vertebra classification using CART suggests that remains of Chinook and sockeye dominate, with steelhead, coho, and cutthroat also present. Remains of Chinook and sockeye had been previously identified by Butler and Baker (2003)

using traditional morphological comparisons, thus the new analysis corroborates and greatly extends the previous work by establishing the presence of a much larger range of species used.

Excluding chum and pink salmon in our more restricted classification exercise requires additional justification. We argue that the rockshelter fish remains represent locally caught fishes rather than those from fishes traded in. Ethnographically, fishing took place along the upper Wenatchee River in summer and fall by local groups of Wenatchi people who occupied winter base camps on the lower river and on the mainstem Columbia (Stevens 2003). The high frequency of salmonid remains (representing all parts of the skeleton), and the site's proximity to rapids where fishes could be easily caught, suggest the site functioned primarily as a fishing camp. If it is accepted that fishes were locally caught, then we turn to historic and contemporary biogeographic salmonid species distributions to justify excluding chum and pink salmon from the classification model. With a few exceptions, these species were confined to the lower reaches of the Columbia River and its tributaries, west of the Cascade crest (Hard et al. 1996; Johnson et al. 1997). Although pink salmon have not been observed in the Wenatchee River basin in recent history, there is some evidence that they have migrated from the ocean to the Snake River Basin, a distance equivalent to that of the confluence of the Wenatchee and Columbia Rivers (Basham and Gilbreath 1978). Counts of upriver-migrating adult salmon moving past the John Day and McNary hydroelectric dams, about 300 miles downstream of the confluence, have infrequently included small numbers of pink salmon in recent years (Fish Passage Center, 1827 NE 44<sup>th</sup> Ave., Suite 240, Portland, OR 97213, [www.fpc.org](http://www.fpc.org)). It is certainly possible that species abundance observed today or noted in historical records varies from that in the deeper past. Study of archaeological salmon records provides an opportunity to empirically establish biogeography of past fish populations and how that may have varied over time in

response to cultural or natural forces (e.g., Chatters et al. 1995; Robinson et al. 2009). We used the rarity of chum and pink salmon from recent records to adjust our classification model and suggest that future applications might also incorporate prior knowledge about species distribution in classification, depending on project goals.

Chum and pink salmon cannot be excluded entirely from the possible fish faunal assemblage at the rockshelters site, but the likelihood of their presence is very small. Given the evidence of chum salmon occurrence primarily in the lower reaches of the Columbia River (Johnson et al. 1997), any appearance of chum in the excavated samples was most likely due to classification error. Furthermore, the models may have had problems differentiating pink and sockeye. The percentage occurrence of pink salmon as assigned by our full set models was about 24% in both upper and lower strata (Table 7), which is high considering the small contribution, if any, pink salmon might have made to the fish faunal assemblage. When we removed chum and pink and refit our classification models, most vertebrae assigned as pink salmon were reassigned as sockeye--the species that most closely matched the vertebral size of pinks in the reference collection (Figure 4) and is one of the dominant species found at the Swiftwater Rockshelters site. DNA analysis would help us to resolve the problems of our models in differentiating pink and sockeye.

Given the aggregate species assignment in CART, being able to distinguish steelhead visually from other species of salmon and trout, especially coho and chum, would be very useful. Gobalet et al. (2004) suggest that vertebrae from steelhead can be distinguished from other salmonids based on size and arrangement of pores on the lateral sides of the centrum. In particular, steelhead vertebrae are thought to have a more robust and “woven” appearance than

other species although Gobalet et al. (2004) note species assignment can be ambiguous. We intend to address this problem of differentiating steelhead from other salmonids in a future study. See Fig. 6 which compares centra from Chinook [A] and steelhead [B].

Cutthroat trout vertebrae were rare in the excavated specimens. Our species classification models were developed using coastal cutthroat trout (*O. clarki clarki*), which were very distinctive in shape compared to the other vertebrae in the modern reference collection. The cutthroat found in the excavated samples most likely were from inland/westslope cutthroat trout (*O. clarki lewisi*; Behnke 2002), which may have been similar enough in size to their coastal relatives for our models to resolve correctly.

Knowledge about salmonid species in site deposits provides several insights that would be completely lost if identifications were left to the genus level. First, the identification of coho salmon vertebrae in the archaeological sample is important because the species is now extinct in the Wenatchee River basin. However, historical records suggest the Wenatchee River spawning population may have numbered in the several thousands, and wild coho were once common in the upper and mid-Columbia region (Mullan 1984; Mullan et al. 1992). Remains excavated from the Swiftwater Rockshelters could be further studied to better understand the genetics of the now extinct population. Comparison of species representation between the lower and upper strata in the rockshelter highlights several patterns, which again would have been obscure without species identifications (Table 7, Fig. 5, see “reduced set”). The lower stratum (occupied 2400-2900 BP) is represented by an even mix of remains identified as Chinook and sockeye (42-46%) and very small quantities of coho and steelhead. In the upper stratum (occupied 1000 BP), sockeye representation increases to 54% while the frequency of Chinook decreases (12%) and the

frequency of steelhead (23%) and coho (10%) increases (Table 7). Changing seasonality in site use may play a role, though the way this would work is complex. Historically, in May and June, runs of Chinook and steelhead migrated past the Swiftwater Rockshelters en route to upriver spawning areas. In August and September, other runs of these species, along with sockeye, migrated upriver (Craig and Suomela 1941). Coho are believed to have migrated into the system later in the fall (Mullan 1984). Given the prominence of sockeye in both strata, late summer/fall occupation is minimally suggested. Schalk (1984) has pointed out that fall run salmon would have been favored for processing for storage (over earlier running fishes) given their overall lower fat content. Butler and Baker's (2003) study of body part representation at the rockshelter suggests fishes represented in both strata were being processed for storage. We are still developing hypotheses to explain the trends in species representation. For now we simply emphasize that without species-level assignments, we would not be aware of any trends at all.

## **5.0 Conclusions**

Our study is important in several ways. Most simply, it provides an additional and rapid, nondestructive tool that can be used to determine species or species group from salmonid vertebrae. Much as species-level identifications from aDNA analysis have allowed researchers to ask much more detailed questions about past human-salmon relationships than allowed by genus-level assignments (e.g., Cannon and Yang 2006), our morphometric approach will facilitate future research with larger sample sizes and thus have the potential for drawing more robust conclusions. Future studies that combine morphometric classification with aDNA analyses will be especially worthwhile. Our quantitative approach also addresses important concerns raised by Driver (1992) and Gobalet (2001) about the need for zooarchaeology to

incorporate more rigorous methods in taxonomic assignments. As we increasingly work to have zooarchaeological research applied to conservation biology and policy debates (e.g., Frazier 2010), the need for greater rigor in analytic protocols and approaches increases in kind.

We suggest two main areas for future work. Our study focused on the seven species of *Oncorhynchus* known for the eastern Pacific; an eighth species, masu (*O. masou*) occupies waters of the western Pacific and inland rivers and streams of east Asia and Siberia (Augerot and Foley 2005). To assist fish zooarchaeology in the western Pacific (e.g., Japan- Matsui 1996; Siberia -Fitzhugh et al. 2004), modern vertebrae from masu salmon could be measured and brought into the classification models. Also, given the difficulty our model had in distinguishing two species groups (steelhead/coho/chum; sockeye/pink), finding additional criteria to distinguish these species would be useful. In this regard, additional work to evaluate criteria proposed by Gobalet et al. (2004:806) to distinguish steelhead vertebrae from other *Oncorhynchus* species would be especially worthwhile. Our morphometric classification and aDNA analysis could be employed for the tests.

## ACKNOWLEDGEMENTS

Drafts of this manuscript were much improved by comments from Jeff Laake, Bryan Wright, Aubrey Cannon, and an anonymous *JAS* reviewer. Salmon carcasses used in our measurements were donated by several hatcheries: South Tacoma Hatchery (chum), Cowlitz Hatchery (cutthroat), Hoodspout Hatchery (pink), Humptulips Hatchery (steelhead) and Chiwawa Hatchery

Ponds (Chinook salmon), Eastbank Hatchery (sockeye) as well as various fishermen (Robert DeLong, Tony Gades, and Pat Gearin). Assistance in measurements was provided by Andrea Gemmer and Sara Fineseth. Bones were photographed by Karna McKinney. Access to bones from the Swiftwater Rockshelters was provided by Kate Valdez, Tribal Historic Preservation Officer, Confederated Tribes and Bands of the Yakama Nation; Camille Pleasants, Tribal Historic Preservation Officer, Confederated Tribes of the Colville Reservation; and Powys Gadd, Heritage Program Leader/Forest Anthropologist, Okanogan-Wenatchee National Forest. Thanks to all.

## LITERATURE CITED

- Adams, P.B., Botsford, L.W., Gobalet, K.W., Leidy, R.A., McEwan, D.R., Moyle, P.B., Smith, J.J., Williams, J.G., Yoshiyama, R.M. 2007. Coho salmon are native south of San Francisco Bay: A reexamination of North American coho salmon's southern range limit. *Fisheries* 32(9), 441-451.
- Augerot, X., Foley, D. N. 2005. *Atlas of Pacific Salmon*. University of California Press, Berkeley.
- Basham, L., Gilbreath, L. 1978. Unusual occurrence of pink salmon (*Oncorhynchus gorbuscha*) in the Snake River of southeastern Washington. *Northwest Science* 52, 32-34.
- Behnke, R.J. 2002. *Trout and salmon of North America*. New York: Free Press. 359 pp.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C. J. 1984. *Classification and regression trees*. Chapman and Hall, New York, NY.
- Breiman, L. 2001. RandomForests. *Machine Learning* 45, 5-32.
- Butler, V.L. 1990. Distinguishing natural from cultural salmonid deposits in Pacific Northwest North America. Ph.D. dissertation, unpublished. University of Washington, Seattle.
- Butler, V.L. 1993. Natural versus cultural salmonid remains: origin of The Dalles roadcut bones, Columbia River, Oregon, U.S.A. *Journal of Archaeological Science* 20, 1-24.
- Butler, V.L., Chatters, J.C. 1994. The role of bone density in structuring prehistoric salmon bone assemblages. *Journal of Archaeological Science* 21, 413-424.
- Butler, V.L., Baker, G. 2003. Appendix E: Fish remains from Swiftwater Rockshelters (45CH433). Pp. E1-E26. In: R. Stevens (Ed.), *Swiftwater Rockshelters (45CH433): Results of data recovery excavations on the Upper Wenatchee River, Chelan County, Washington*. Eastern Washington University Reports in Archaeology and History 100-120, Archaeological and Historical Services, Eastern Washington University, Cheney, WA. 128 p.
- Butler, V.L., Campbell, S.K. 2004. Resource intensification and resource depression in the Pacific Northwest of North America: a zooarchaeological review. *Journal of World Prehistory* 18(4), 327-405.
- Butler, V.L., Misarti, N., Yang, D.Y., Speller, C.F., Connolly, T.J., Jenkins, D.L. Unpublished manuscript. Using archaeological fish remains to determine species of salmon and trout in the Upper Klamath Basin before hydro-development.
- Cannon, A. 1988. Radiographic age determination of Pacific salmon: species and seasonal inferences. *Journal of Field Archaeology* 15:103-108.



- Cannon, A. 2000. Assessing variability in Northwest coast salmon and herring fisheries: Bucket-auger sampling of shell midden sites on the central coast of British Columbia. *Journal of Archaeological Science* 27(8), 725-737.
- Cannon, A., Yang, D.Y. 2006. Early storage and sedentism on the Pacific Northwest coast: Ancient DNA analysis of salmon remains from Namu, British Columbia. *American Antiquity* 71, 123-140.
- Casteel, R.W. 1974. Identification of the species of Pacific salmon (genus *Oncorhynchus*) native to North America based upon otoliths. *Copeia*, 1974(2):305-311.
- Casteel, R.W. 1976. Fish remains in archaeology and paleo-environmental studies. Academic Press, New York.
- Chatters, J. C., Butler, V.L., Scott, M.J., Anderson, D.M., Neitzel, D.A. 1995. A paleoscience approach to estimating the effects of climatic warming on salmonid fisheries of the Columbia River basin. Pp. 489-496 *In* R.J. Beamish (ed.). Climate change and northern fish populations. Canadian Special Publication of Fisheries and Aquatic Sciences 121.
- Craig, J.A., Suomela, A.J. 1941. Time of appearance of the runs of salmon and steelhead trout native to the Wenatchee, Entiat, Methow, and Okanogan Rivers. Unpublished Manuscript reprinted in U.S.F.W Monograph 1, 1992: Production and Habitat of Salmonids in Mid-Columbia River Tributary Streams by J.W. Mullan, K.R. Williams, G. Rhodus, T.W. Hillman, and J.D. McIntyre.
- Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J. 2007. Random forests for classification in ecology. *Ecology* 88, 2783-2792.
- Driver, J.C. 1992. Identification, classification and zooarchaeology. *Circaea* 9, 35-47.
- Feldesman, M.R. 2002. Classification trees as an alternative to linear discriminant analysis, *American Journal of Physical Anthropology* 119, 257-275.
- Fitzhugh, B., Moore, S., Lockwood, C., Boone, C. 2004. Archaeological paleobiogeography in the Russian Far East: The Kuril Islands and Sakhalin in comparative perspective. *Asian Perspectives* 43(1), 92-122.
- Frazier, J. G. 2010. The call of the wild. In: Rebecca M. Dean (Ed.), The archaeology of anthropogenic environments. Center for Archaeological Investigations, Occasional Paper No. 37, Southern Illinois University, Carbondale, pp. 341-369.
- Gobalet, K.W. 2001. A critique of faunal analysis: inconsistency among experts in blind tests. *Journal of Archaeological Science* 28(4), 377-386.

- Gobalet, K.W., Schulz, P.D., Wake, T.A., Siefkin, N. 2004. Archaeological perspectives on Native American fisheries of California, with emphasis on steelhead and salmon. Transactions of the American Fisheries Society 133, 801–833.
- Gorschkov, S.A., Dorofeyeva, Y.A., Klyukanov, V.A., Kulikova, N.I. 1979. Osteological characteristics of the Pacific salmon in the genus *Oncorhynchus*. Journal of Ichthyology 19:1-19.
- Hard, J.J., Kope, R.G., Grant, W.S., Waknitz, F.W., Parker, L.T., Waples, R.S. 1996. Status review of pink salmon from Washington, Oregon, and California. U. S. Department of Commerce, NOAA Tech. Memo. NMFS-NWFSC-25. Seattle, WA. 131 p.
- Johnson, O. W., Grant, W.S., Kope, R.G., Neely, K., Waknitz, F.W., Waples, R.S. 1997. Status review of chum salmon from Washington, Oregon, and California. U. S. Department of Commerce, NOAA Tech. Memo. NMFS-NWFSC-32. 280 p.
- Karels, T.J., Bryant, A.A., Hik, D.S. 2004. Comparison of discriminant function and classification tree analyses for age classification of marmots. Oikos 105, 575-587.
- Liaw, A., Wiener, M. 2002. Classification and regression by randomForest. R News 2(3), 18-22.
- Lyman, R.L. 2003. Appendix C: Mammal, reptile, and bird remains from 45CH433. Pp. C1-C24. In: R. Stevens (Ed.), Swiftwater Rockshelters (45CH433): Results of data recovery excavations on the Upper Wenatchee River, Chelan County, Washington. Eastern Washington University Reports in Archaeology and History 100-120, Archaeological and Historical Services, Eastern Washington University, Cheney, WA. 128 p.
- Matsui, A. 1996. Archaeological investigations of anadromous salmonid fishing in Japan. World Archaeology 27(3), 444-460.
- Morales, A. 1984. A study on the representativity and taxonomy of the fish faunas from two Mousterian sites on northern Spain with special reference to the trout (*Salmon trutta* L., 1758). In N. Desse-Berset (Ed.), 2nd Fish Osteoarchaeology Meeting. Notes et Monographies Techniques. 16 Centre de Recherches Archeologiques, Paris, pp. 41-59.
- Mullan, J.W. 1984. Overview of artificial and natural propagation of coho salmon (*Oncorhynchus kisutch*) on the Mid-Columbia River. U.S. Fish and Wildlife Service, Rep. No. FRI/FAO-84-4, Leavenworth, WA.
- Mullan, J.W., Williams, K.R., Rhodus, G., Hillman, T.W., McIntyre, J.D. 1992. Production and habitat of salmonids in mid-Columbia River tributary streams. U.S. Fish and Wildlife Service, Monograph I.
- Orchard, T.J. Szpak, P. Unpublished results. Identification of salmon species from archaeological remains on the Northwest Coast. In: Moss, M., Cannon, A. (Eds.), The Archaeology of North Pacific Fisheries. University of Alaska Press.

- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available online: URL <http://www.R-project.org> (accessed May 2008).
- Ripley, B. 1996. Pattern recognition and neural networks. Cambridge University Press, New York, NY.
- Robinson, B.S., Jacobson, G.L., Yates, M.G., Spiess A.E., Cowie, E.R. 2009. Atlantic salmon, archaeology and climate change in New England. *Journal of Archaeological Science* 36, 2184-2191.
- Schalk, R.F. 1984. The Columbia Plateau salmon fisheries: archaeological evidence for intensification. Paper presented at the Annual Meeting of the Society for American Archaeology, Portland, OR.
- Speller, C. F., Yang, D.Y., Hayden, B. 2005. Ancient DNA investigation of prehistoric salmon resource utilization at Keatley Creek, British Columbia, Canada. *Journal of Archaeological Science* 32, 1378-1389
- Stevens, R. 2003. Swiftwater Rockshelters (45CH433): Results of data recovery excavations on the Upper Wenatchee River, Chelan County, Washington. Eastern Washington University Reports in Archaeology and History 100-120, Archaeological and Historical Services, Eastern Washington University, Cheney, WA. 128 p.
- Therneau, T.M., Atkinson, E.J. 1997. An introduction to recursive partitioning using the RPART routines. Technical Report No. 61. Department of Health Sciences Research, Section of Biostatistics, Mayo Clinic, Rochester, MN. Available online: [www.mayo.edu/hsr/techrpt/61.pdf](http://www.mayo.edu/hsr/techrpt/61.pdf) (accessed May 2008).
- Therneau, T.M., Atkinson, B. 2008. rpart: recursive partitioning. R package version 3.1-41. R port by Brian Ripley.
- Venables, W.N., Ripley, B.D. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.
- Weinand, D.C. A study of parametric versus non-parametric methods for predicting paleohabitat from Southeast Asian Bovid astragali. *Journal of Archaeological Science* 34, 1774-1783.
- Yang, D.Y., A. Cannon and S.R. Saunders. 2004. DNA species identification of archaeological salmon bone from the Pacific Northwest coast of North America. *Journal of Archaeological Science* 31, 619-631.

# TABLES

Table 1: Pacific salmonid species sample specimens used for this study.

Species	<i>n</i>	Total length		Collection areas (Washington State)
			range	
Chinook	14	73 cm - 95 cm	Neah Bay Troll fishery, Tulalip Tribal fishery, Nisqually Reach, University of Washington Hatchery, Columbia River (Drano Lake)	
Chum	10	66 cm - 80 cm	South Tacoma Hatchery	
Coho	10	51 cm - 74 cm	Sekiu River, Sol Duc River, Neah Bay Tribal fishery, Snohomish River	
Pink	10	51 cm - 62 cm	Sekiu River, Hoodspout Hatchery	
Sockeye	11	48 cm - 60 cm	Neah Bay, Eastbank Hatchery	
Steelhead	14	63 cm - 84 cm	Humptulips Hatchery, Neah Bay, Skykomish River	
Cutthroat	10	27 cm - 35 cm	Cowlitz Hatchery	

Table 2: Characteristics of salmonid vertebrae from the modern collection used in this study. Mean lengths and heights are in millimeters. Type II vertebrae lack fused processes. Type III vertebrae have fused processes (see Fig. 1).

Species	Type II					Type III				
	n	Mean	Variance	Standard error	Coefficient of variation	n	Mean	Variance	Standard error	Coefficient of variation
<b>Length</b>										
Chinook salmon	356	6.93	1.023	0.054	0.146	484	7.43	0.771	0.040	0.118
chum	310	6.87	0.311	0.032	0.081	262	7.29	0.188	0.027	0.059
coho	248	6.53	0.354	0.038	0.091	318	6.83	0.414	0.036	0.094
pink	269	4.69	0.151	0.024	0.083	288	5.38	0.125	0.021	0.066
sockeye	330	4.82	0.244	0.027	0.102	294	5.53	0.180	0.025	0.077
steelhead	353	6.47	0.951	0.052	0.151	375	7.44	0.828	0.047	0.122
cutthroat	284	3.30	0.134	0.022	0.111	230	3.60	0.079	0.019	0.078
<b>Height</b>										
Chinook salmon	356	10.21	1.460	0.064	0.118	484	10.94	1.354	0.053	0.106
chum	310	8.17	0.644	0.046	0.098	262	9.15	0.267	0.032	0.056
coho	248	8.44	0.505	0.045	0.084	318	8.94	0.569	0.042	0.084
pink	269	6.30	0.614	0.048	0.124	288	6.89	0.396	0.037	0.091
sockeye	330	5.97	0.395	0.035	0.105	294	6.49	0.371	0.036	0.094
steelhead	353	7.92	1.161	0.057	0.136	375	8.74	1.041	0.053	0.117
cutthroat	284	3.59	0.124	0.021	0.098	230	3.85	0.094	0.020	0.079
<b>Length/Height ratio</b>										
Chinook salmon	356	0.68	0.0020	0.0024	0.066	484	0.68	0.0022	0.0021	0.069

Species	Type II					Type III				
	n	Mean	Variance	Standard error	Coefficient of variation	n	Mean	Variance	Standard error	Coefficient of variation
chum	310	0.84	0.0020	0.0025	0.053	262	0.80	0.0010	0.0020	0.040
coho	248	0.77	0.0012	0.0022	0.045	318	0.76	0.0016	0.0022	0.052
pink	269	0.75	0.0020	0.0028	0.060	288	0.79	0.0031	0.0033	0.071
sockeye	330	0.81	0.0014	0.0021	0.046	294	0.85	0.0020	0.0026	0.053
steelhead	353	0.82	0.0017	0.0022	0.050	375	0.85	0.0018	0.0022	0.050
cutthroat	284	0.92	0.0027	0.0031	0.056	230	0.94	0.0014	0.0024	0.040

Table 3: Within and overall training data success rates of salmonid vertebra for three classification techniques, Linear Discriminant Analysis (LDA), classification and regression trees (CART), and randomForest using vertebral morphological features. Roman numerals correspond to vertebral types included in the classification procedures. Numbers in **bold** indicate the highest (or ties for the highest) classification method success rates within vertebral type groups (II, III, or II and III combined).

Species	LDA			CART			RandomForest		
	II	III	II & III	II	III	II & III	II	III	II & III
Overall success rate	0.71	0.74	0.68	<b>0.77</b>	<b>0.78</b>	<b>0.73</b>	0.75	0.76	0.70
Chinook	0.90	0.87	0.89	<b>0.95</b>	<b>0.96</b>	<b>0.95</b>	<b>0.95</b>	<b>0.96</b>	<b>0.95</b>
Chum	0.64	0.29	0.14	<b>0.73</b>	0.36	<b>0.73</b>	0.63	<b>0.48</b>	0.48
Coho	0.58	<b>0.65</b>	<b>0.63</b>	0.60	0.58	0.60	<b>0.64</b>	0.54	0.54
Pink	0.66	0.66	0.61	<b>0.68</b>	<b>0.72</b>	<b>0.68</b>	0.65	0.71	0.57
Sockeye	0.80	0.76	0.77	<b>0.81</b>	<b>0.88</b>	<b>0.81</b>	0.73	0.74	0.70
Steelhead	0.48	<b>0.83</b>	<b>0.69</b>	0.57	<b>0.83</b>	0.57	<b>0.60</b>	0.79	0.60
Cutthroat	0.91	0.97	0.94	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>

Table 4: Within and overall success rates of test data using the fits from the model training set. Sample sizes included five vertebrae of each vertebral type per species. Numbers in **bold** indicate the highest (or ties for the highest) classification method success rates within vertebral type groups (II, III, or II and III combined).

Species	LDA			CART			randomForest		
	II	III	II & III	II	III	II & III	II	III	II & III
Chinook	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>
Chum	<b>0.6</b>	0.2	0.2	0.4	0.2	<b>0.5</b>	<b>0.6</b>	<b>0.4</b>	0.4
Coho	<b>0.4</b>	0.8	<b>0.8</b>	<b>0.4</b>	0.6	0.5	<b>0.4</b>	<b>1.0</b>	0.6
Pink	0.8	<b>0.8</b>	<b>0.9</b>	<b>1.0</b>	<b>0.8</b>	<b>0.9</b>	0.6	<b>0.8</b>	0.3
Sockeye	<b>1.0</b>	0.6	<b>0.8</b>	<b>1.0</b>	<b>1.0</b>	<b>0.8</b>	<b>1.0</b>	0.6	<b>0.8</b>
Steelhead	<b>0.4</b>	<b>0.8</b>	<b>0.6</b>	<b>0.4</b>	0.6	0.3	<b>0.4</b>	0.6	0.4
Cutthroat	0.6	0.8	0.7	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>



Table 5: Within and overall training data success rates of salmonid vertebra for three classification techniques on grouped species, Linear Discriminant Analysis (LDA), classification and regression trees (CART), and randomForest using vertebral morphological features. Roman numerals correspond to vertebral types included in the classification procedures. Numbers in **bold** indicate the highest (or ties for the highest) classification method success rates within vertebral type groups (II, III, or II & III combined).

Species	LDA			CART			randomForest		
	II	III	II & III	II	III	II & III	II	III	II & III
Overall success rate	0.91	0.91	0.89	<b>0.94</b>	<b>0.96</b>	<b>0.92</b>	0.93	0.95	0.91
Chinook	0.91	0.88	0.90	<b>0.96</b>	<b>0.96</b>	0.94	0.95	<b>0.96</b>	<b>0.95</b>
Chum, coho, steelhead	0.67	0.92	0.87	0.91	<b>0.95</b>	<b>0.91</b>	<b>0.92</b>	0.93	0.89
Cutthroat	0.92	0.97	0.94	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
Pink, sockeye	<b>0.94</b>	0.91	<b>0.89</b>	<b>0.94</b>	<b>0.96</b>	<b>0.89</b>	0.91	0.95	<b>0.89</b>

Table 6: Within and overall success rates of test data using the fits from the model training set.

Sample sizes included five vertebrae of each vertebral type per species or species group.

Numbers in **bold** indicate the highest (or ties for the highest) classification method success rates within vertebral type groups (II, III, or II and III combined).

Species	LDA			CART			randomForest		
	II	III	II & III	II	III	II & III	II	III	II & III
Chinook	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>
Chum, coho, steelhead	0.8	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>
Cutthroat	<b>1.0</b>	0.8	0.9	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>
Pink, sockeye	<b>1.0</b>	0.8	<b>0.9</b>	<b>1.0</b>	<b>1.0</b>	<b>0.9</b>	<b>1.0</b>	<b>1.0</b>	<b>0.9</b>

Table 7: Percent composition of salmonid species or species group from the Swiftwater Rockshelters excavation by strata, identified using classification models for the full set of seven Pacific salmonids in the classification models, a reduced set which excluded chum and pink salmon from the model, and results based on species groups. Upper stratum = 1040 +/-50 ybp and lower stratum = 2420-2900 +/- 60 ybp.

Stratum	n	Percent salmonid species						
		Chinook	Chum	Coho	Steelhead	Cutthroat	Pink	Sockeye
Full set								
Upper Stratum	145	12.41	0.69	3.45	17.93	0.69	24.14	40.69
Lower Stratum	164	42.07	0	1.22	5.49	0	23.17	28.05
Reduced set (chum and pink removed)								
Upper Stratum	145	12.41	---	9.66	22.76	0.69	---	54.48
Lower Stratum	164	42.07	---	4.27	6.71	0	---	46.95
Species groupings	n	Chum, coho and steelhead combined			Cutthroat	Pink and sockeye combined		
		Chinook						
Upper Stratum	145	13.10		26.21	0.69		60.00	
Lower Stratum	164	42.07		7.93	0		50.00	

## FIGURE CAPTIONS

Fig. 1. Photograph of salmonid vertebrae: Type II, showing centra with unfused processes (A1, A2, A3); and Type III, showing fused dorsal/haemal spines (B).

Fig. 2: Photograph showing caliper position and vertebrae measures obtained for classification models: height (A), and length (B).

Fig. 3: Map of east-central Washington State, showing location of Lake Wenatchee, Columbia River and Snake River.

Fig. 4: Vertebral height measurements from seven west coast Pacific salmonid species (Chinook, chum, coho, cutthroat, pink, sockeye, and steelhead) showing overlap in measurements for five species in two groupings. Bars denote  $\pm 1$  standard error.

Fig. 5: Frequency of *Oncorhynchus* species in the upper (a) and lower (b) strata of the Swiftwater Rockshelters site, based on vertebrae and using CART classification. Left side (dark bars) depicts frequency with all seven Pacific salmonid species included. Right side (light bars) depicts frequency with the reduced set (chum and pink salmon removed).

Fig. 6: Photograph comparing bone texture of Chinook (A) and steelhead (B).

Figure 1

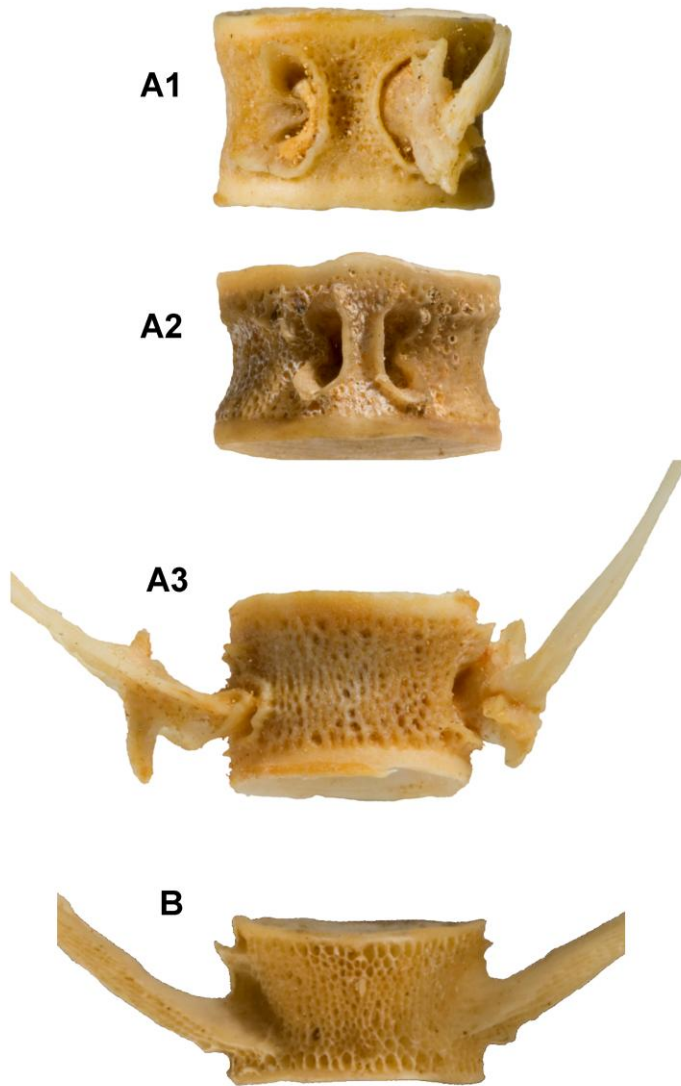
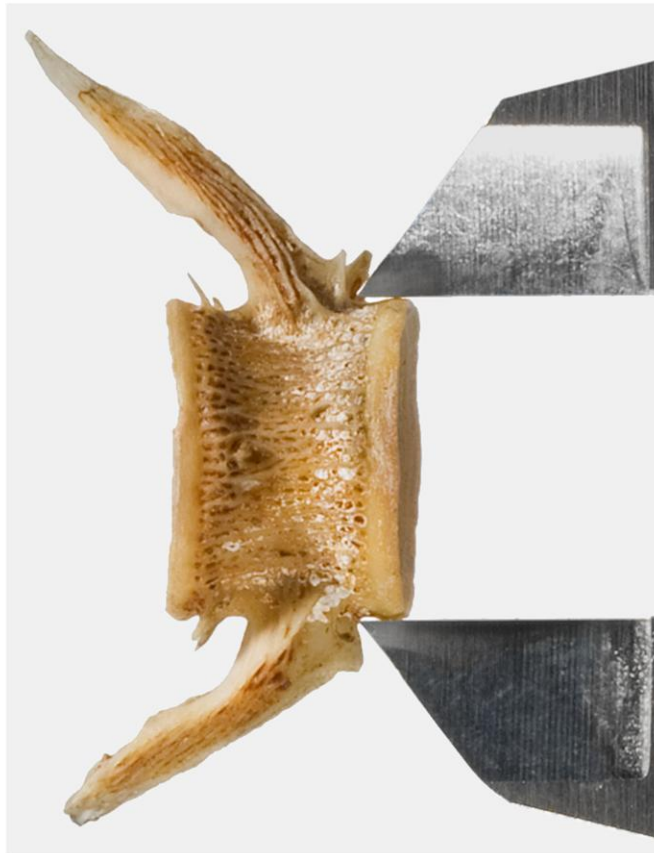


Figure 2

A



B

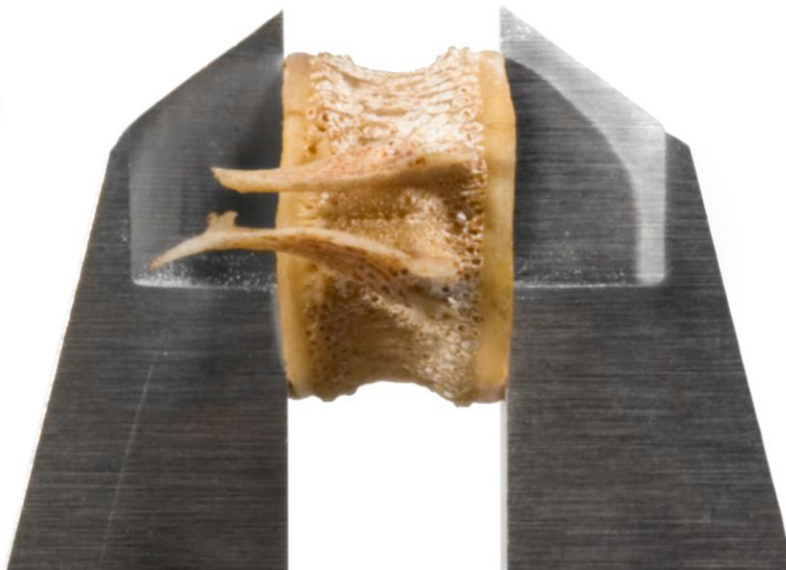


Figure 3

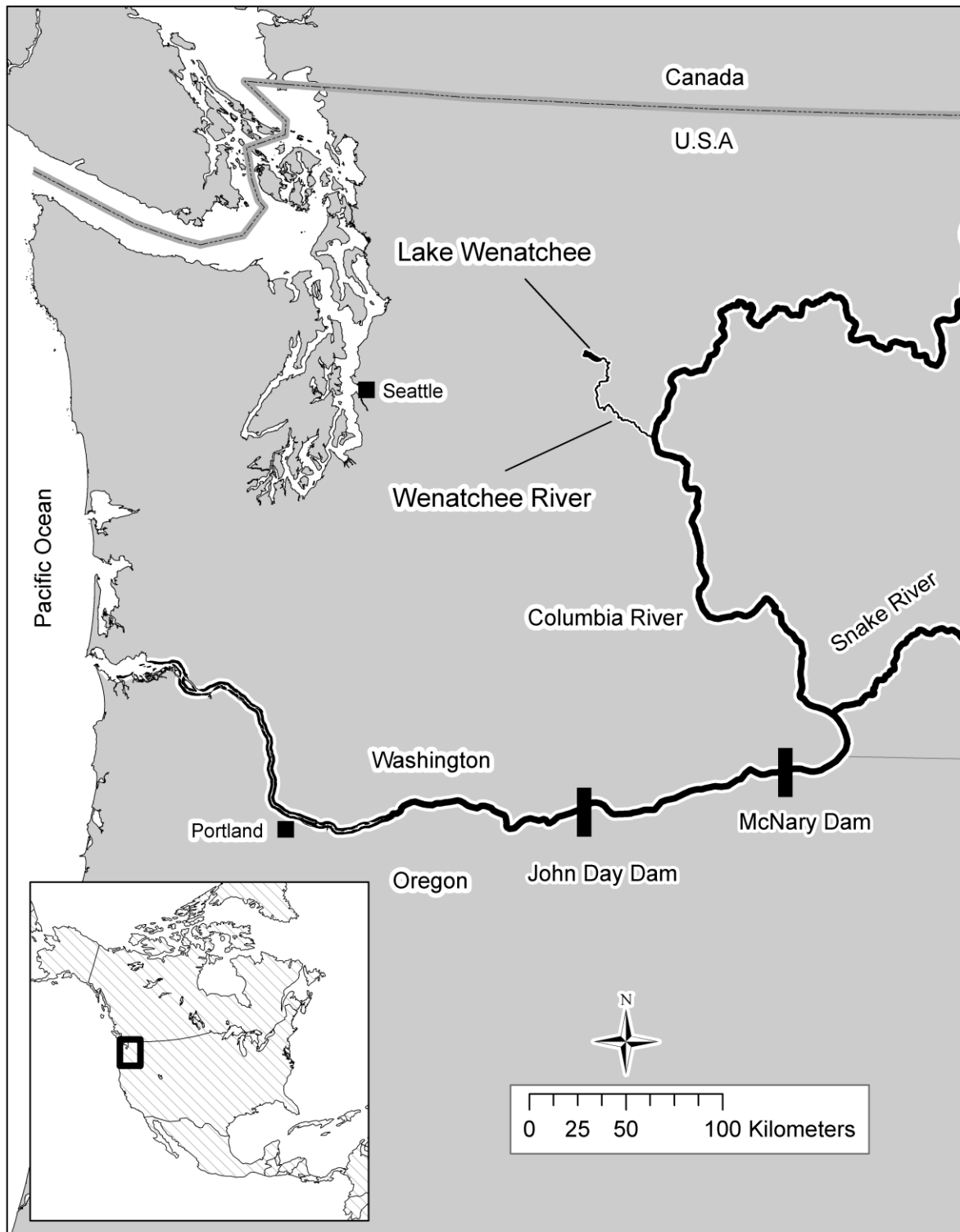


Figure 4

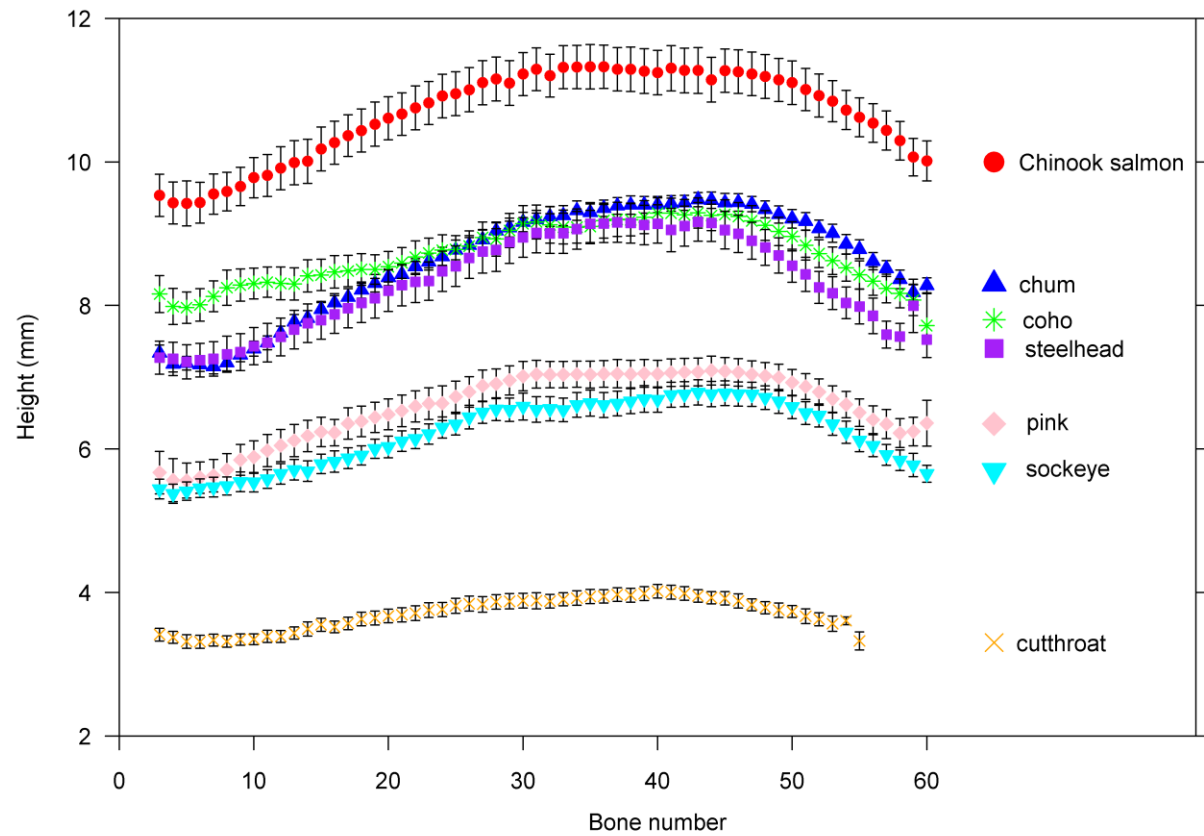




Figure 5

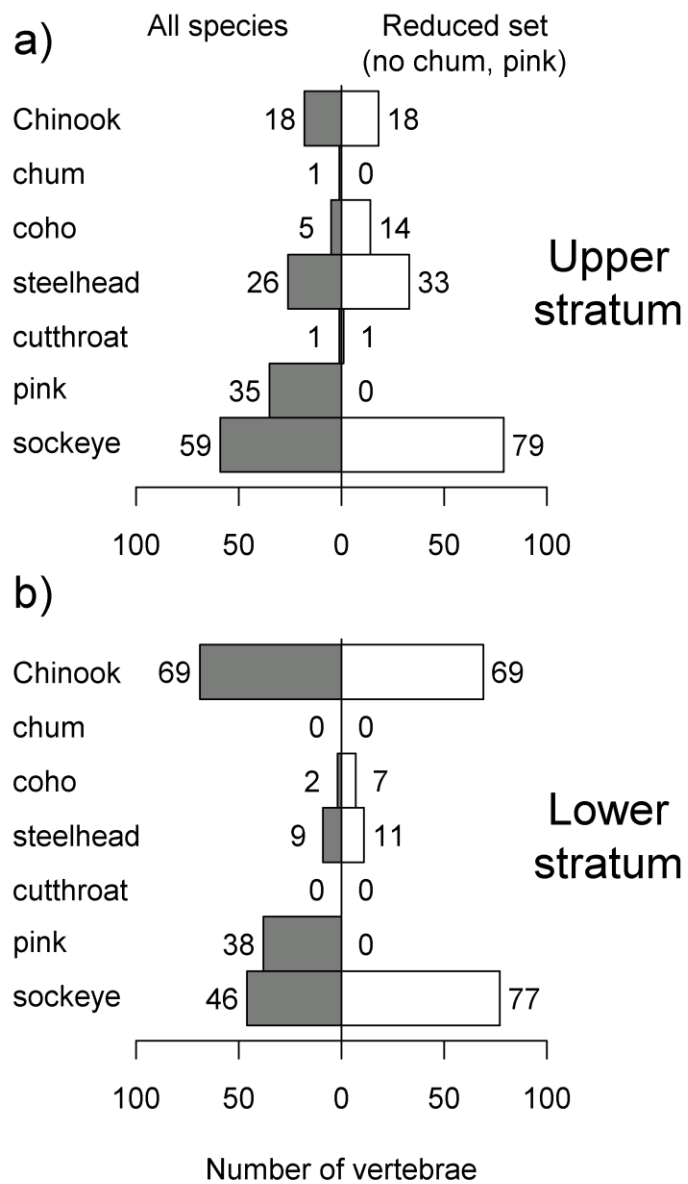


Figure 6

